GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 21, 2002, 10:15:23 ; Search time 30.33 Seconds
(without alignments)
1102.315 Million cell updates/sec

Title: US-09-758-017A-2
Perfect score: 1603
Sequence: 1 MLFKLGLCQRCISSNRVLPG.....SKANGLLKLSGTEPINWRAL 301

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1596	99.6	301	22	AAU04939	Atlantic cod heat-
2	1421	88.6	301	22	AAU04940	Atlantic cod heat-
3	1030.5	64.3	292	22	AAG74939	Human colon cancer
4	1018.5	63.5	313	18	AAW21814	Uracil DNA glycosy
5	1015.5	63.3	313	18	AAW21823	Mutant uracil DNA
6	1014.5	63.3	313	18	AAW21822	Mutant uracil DNA
7	1014.5	63.3	313	18	AAW21820	Mutant uracil DNA
8	1013.5	63.2	313	18	AAW21821	Mutant uracil DNA
9	1013.5	63.2	313	18	AAW21815	Cytosine DNA glyco
10	1012.5	63.2	313	18	AAW21824	Mutant uracil DNA
11	1011.5	63.1	313	18	AAW21819	Mutant uracil DNA
12	1009.5	63.0	313	18	AAW21816	Thymine DNA glycos
13	1007.5	62.9	313	18	AAW21825	Mutant uracil DNA
14	681.5	42.5	229	22	AAU69754	Escherichia coli u
15	635.5	39.6	219	22	AAB88498	Haemophilus influe
16	598.5	37.3	321	21	AAG14409	Arabidopsis thalia
17	598.5	37.3	330	21	AAG14408	Arabidopsis thalia
18	591.5	36.9	227	20	AAY01335	B. pallidus uracil
19	539.5	33.7	236	20	AAY35405	Chlamydia pneumoni
20	539	33.6	229	20	AAY37595	Amino acid sequenc
21	520	32.4	216	22	AAG81666	S. epidermidis ope
22	444.5	27.7	255	21	AAB53166	Macaca mulatta rha
23	423	26.4	334	19	AAW72171	HSV-2 strain SB5 C
24	423	26.4	334	19	AAW72090	HSV-2 strain SB5 C
25	423	26.4	334	19	AAW72005	HSV-2 strain SB5 C
26	415	25.9	293	19	AAW80437	Feline herpesvirus
27	384	24.0	233	19	AAW98728	H. pylori GHPO 902
28	366.5	22.9	254	22	AAG91207	C glutamicum prote
29	345	21.5	264	22	AAU55037	Propionibacterium
30	303	18.9	995	22	ABG11717	Novel human diagno
31	275	17.2	72	21	AAB44117	Human cancer assoc
32	125	7.8	29	20	AAY42895	Human wild-type ur
33	116	7.2	29	20	AAY50001	Human mutant uraci
34	112	7.0	29	20	AAY50002	Human mutant uraci
35	104	6.5	1069	22	ABB63364	Drosophila melanog
36	90.5	5.6	506	22	ABB72057	Drosophila melanog
37	89.5	5.6	931	22	AAM39815	Human polypeptide
38	89	5.6	752	15	AAR51701	AML1-MTG8 fusion.
39	87.5	5.5	1005	22	ABG11447	Novel human diagno
40	87	5.4	1362	22	AAU38416	Salmonella typhi c
41	86	5.4	498	21	AAB13559	Streptomyces globi
42	84.5	5.3	797	17	AAW05290	Triticum tauschii
43	84.5	5.3	2271	22	ABB65616	Drosophila melanog
44	83	5.2	29	20	AAY42896	Yeast uracil DNA g
45	82.5	5.1	547	10	AAP94045	t-plastin derived

ALIGNMENTS

RESULT 1
AAU04939
ID AAU04939 standard; Protein; 301 AA.
XX
AC AAU04939;
XX
DT 24-OCT-2001 (first entry)
XX
DE Atlantic cod heat-labile uracil-DNA glycosylase, UNG #1.
XX
KW Atlantic cod; heat-labile uracil-DNA glycolsylase; UNG; UDG;
KW PCR control; LCR control; ligase chain reaction; carry-over prevention.
XX
OS Gadus morhua.
XX
FH Key Location/Qualifiers
FT Misc-difference 17
FT /note= "Encoded by GTY"
XX
PN WO200151623-A1.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-NO00008.
XX
PR 12-JAN-2000; 2000NO-0000163.
PR 27-OCT-2000; 2000NO-0005428.
XX
PA (BIOT-) BIOTEC ASA.
XX
PI Lanes O, Willasen NP, Guddal PH, Gjellesvik DR;
XX
DR WPI; 2001-451854/48.
DR N-PSDB; AAS09498.

XX
PT New cod liver uracil-DNA glycosylase enzyme, useful in monitoring or
PT controlling a reaction system multiplying DNA sequences or in
PT carry-over prevention procedures -
XX
PS Claim 2; Page 52-54; 59pp; English.
XX
CC The sequence is an Atlantic cod heat-labile uracil-DNA glycosylase,
CC (UNG/UDG). The enzyme is useful in monitoring and/or controlling a
CC reaction system multiplying DNA sequences, e.g. PCR (polymerase chain
CC reaction) or LCR (ligase chain reaction). The enzyme is also useful in
CC carry-over prevention procedures.
XX
SQ Sequence 301 AA;

Query Match 99.6%; Score 1596; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 9e-156;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MLFKLGLCQRCISSNRVLPGLLIPQTLCFSKLMKITPKKLRSSNVEQKTSSPQLSVEQLE 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mlfklglcqrcissnrvlpgllipqtlcfsklmkitpkklrssnveqktsspqlsveqle 60

Qy      61 RMAKNKKAALDKIRAKATPAGFGETWRRELAAEFEKPYFKQLMSFVADERSRHTVYPPAD 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 rmaknkkaaldkirakatpagfgetwrrelaaefekpyfkqlmsfvadersrhtvyppad 120

Qy     121 QVYSSTEMCDIQDVKVVILGQDPYHGPNQAHGLCFSVQKPVPPPPSLVNIYKELCTDIDG 180
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 qvyswtemcdiqdvkvvilgqdpyhgpnqahglcfsvqkpvppppslvniykelctdidg 180

Qy     181 FKHPGHGDLSGWAKQGVLLLNAVLTVRAHQANSHKDRGWETFTDAVIKWLSVNREGVVFL 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 fkhpghgdlsgwakqgvlllnavltvrahqanshkdrgwetftdavikwlsvnregvvfl 240

Qy     241 LWGSYAHKKGATIDRKRHHVLQAVHPSPLSAHRGFLGCKHFSKANGLLKLSGTEPINWRA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 lwgsyahkkgatidrkrhhvlqavhpsplsahrgflgckhfskangllklsgtepinwra 300

Qy     301 L 301
           |
Db     301 l 301
```

RESULT 2
AAU04940
ID AAU04940 standard; Protein; 301 AA.
XX
AC AAU04940;
XX
DT 24-OCT-2001 (first entry)
XX
DE Atlantic cod heat-labile uracil-DNA glycosylase, UNG #2.
XX
KW Atlantic cod; heat-labile uracil-DNA glycolsylase; UNG; UDG;
KW PCR control; LCR control; ligase chain reaction; carry-over prevention.
XX
OS Gadus morhua.
XX
PN WO200151623-A1.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-NO00008.
XX
PR 12-JAN-2000; 2000NO-0000163.
PR 27-OCT-2000; 2000NO-0005428.
XX
PA (BIOT-) BIOTEC ASA.
XX
PI Lanes O, Willasen NP, Guddal PH, Gjellesvik DR;
XX
DR WPI; 2001-451854/48.
DR N-PSDB; AAS09499.
XX
PT New cod liver uracil-DNA glycosylase enzyme, useful in monitoring or
PT controlling a reaction system multiplying DNA sequences or in
PT carry-over prevention procedures -
XX
PS Claim 2; Page 54-56; 59pp; English.
XX
CC The sequence is an Atlantic cod heat-labile uracil-DNA glycosylase,
CC (UNG/UDG). The enzyme is useful in monitoring and/or controlling a
CC reaction system multiplying DNA sequences, e.g. PCR (polymerase chain
CC reaction) or LCR (ligase chain reaction). The enzyme is also useful in
CC carry-over prevention procedures.
XX
SQ Sequence 301 AA;

Query Match 88.6%; Score 1421; DB 22; Length 301;
Best Local Similarity 98.9%; Pred. No. 9.4e-138;
Matches 266; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      33 MKITPKKLRSSNVEQKTSSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 vqitpkklrssnveqktsspqlsveqlermaknkkaaldkirakatpagfgetwrrelaa 92

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      93 efekpyfkqlmsfvadersrhtvyppadqvyswtemcdiqdvkvvilgqdpyhgpnqahg 152

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 lcfsvqkpvppppslvniykelctdidgfkhpghgdlsgwakqgvlllnavltvrahqan 212

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 shkdrgwetftdavikwlsvnregvvfllwgsyahkkgatidrkrhhvlqavhpsplsah 272

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           |||||||||||||||||||||||||||||
Db     273 rgflgckhfskangllklsgtepinwral 301
```

RESULT 3
AAG74939
ID AAG74939 standard; Protein; 292 AA.
XX
AC AAG74939;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:5703.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 12.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH34344.

```
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 11; Page 7239-7240; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence   292 AA;


  Query Match             64.3%;  Score 1030.5;  DB 22;  Length 292;
  Best Local Similarity   69.0%;  Pred. No. 1.5e-97;
  Matches  191;  Conservative   31;  Mismatches   52;  Indels    3;  Gaps    2;

Qy      27 LCFSKLMKITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGE 84
           ||   |  |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||
Db      17 lcgdhlqaipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfge 75

Qy      85 TWRRELAAEFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPY 144
           :|:: |: || |||| :|| |||:||  :|||||  ||:: |:||||:|||||||||||
Db      76 swkkhlsgefgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpy 135

Qy     145 HGPNQAHGLCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVL 204
           ||||||||||||||:|||||||| |||||| |||: | ||||||||||||||||||||||
Db     136 hgpnqahglcfsvqrpvppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavl 195

Qy     205 TVRAHQANSHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAV 264
           ||||||||||:|||| |||||: ||: |  |:|||||||| |||: ||||||||||
Db     196 tvrahqanshkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqta 255

Qy     265 HPSPLSAHRGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           |||||| :||| ||:|||| | ||: || :||:|: |
Db     256 hpsplsvyrgffgcrhfsktnellqksgkkpidwkel 292


RESULT   4
AAW21814
ID   AAW21814 standard; Protein; 313 AA.
XX
AC   AAW21814;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Uracil DNA glycosylase UNG2.
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis; DNA sequencing;
KW   DNA modification; cell killing.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..44
FT                   /label= Nuclear_localisation_peptide
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
DR   N-PSDB; AAT73564.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Claim 1; Page 47-48; 60pp; English.
XX
CC   A new form of uracil DNA glycosylase (AAW21814), designated UNG2, is
CC   capable of releasing uracil bases from single and/or double
CC   stranded DNA.  Its amino acid sequence was deduced from a cDNA
CC   clone (AAT73564) obtd. from a human NT2 neuronal precursor cell
CC   library.  UNG2 represents an alternatively spliced form of UNG that
CC   includes a 44-amino acid presequence (see also AAW21818) that is not
CC   essential for catalytic activity.  UNG2 can be mutated to provide
CC   cytosine or thymine DNA glycosylases (see also AAW21815-16).
CC   Recombinant DNA glycosylases can be expressed in host cells for use
CC   in mutagenesis, to remove contaminating DNA prior to PCR, in DNA
CC   modification and in cell killing.
XX
SQ   Sequence   313 AA;


  Query Match             63.5%;  Score 1018.5;  DB 18;  Length 313;
  Best Local Similarity   69.9%;  Pred. No. 2.8e-96;
  Matches  188;  Conservative   31;  Mismatches   47;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||||||||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpyhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| ||||||: ||: |  |:|||||||| |||: |||||||||| |||||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT   5
AAW21823
ID   AAW21823 standard; Protein; 313 AA.
XX
AC   AAW21823;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (S169A).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
```

```
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 178
FT                   /note= "site of Ser-178 to Ala substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;


  Query Match             63.3%;  Score 1015.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 5.7e-96;
  Matches  187;  Conservative   32;  Mismatches   47;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||||||||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpyhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||:| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvppppaleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| ||||||: ||: |  |:|||||||| |||: |||||||||  |||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT   6
AAW21822
ID   AAW21822 standard; Protein; 313 AA.
XX
AC   AAW21822;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (Y147F).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 156
FT                   /note= "site of Tyr-147 to Phe substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;


  Query Match             63.3%;  Score 1014.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 7.2e-96;
  Matches  187;  Conservative   32;  Mismatches   47;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||||||||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpfhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| ||||||: ||: |  |:|||||||| |||: |||||||||  |||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT   7
```

```
AAW21820
ID   AAW21820 standard; Protein; 313 AA.
XX
AC   AAW21820;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (D145E).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 154
FT                   /note= "site of Asp-145 to Glu substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;

  Query Match             63.3%;  Score 1014.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 7.2e-96;
  Matches  187;  Conservative   32;  Mismatches   47;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||:|||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqepyhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvpppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| |||||: ||: |  |:|||||||| |||: |||||||||  |||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
```

```
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT   8
AAW21821
ID   AAW21821 standard; Protein; 313 AA.
XX
AC   AAW21821;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (D145N).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 154
FT                   /note= "site of Asp-145 to Asn substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;

  Query Match             63.2%;  Score 1013.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 9.1e-96;
  Matches  187;  Conservative   32;  Mismatches   47;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||:|||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqepyhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvpppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224
```

```
Qy    213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
          |||:|||| ||||| ||: | |:|||||||| ||| |||||||||  |||||| :
Db    225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy    273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
          ||| ||:|||| | ||: || :||:|: |
Db    285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  9
AAW21815
ID   AAW21815 standard; Protein; 313 AA.
XX
AC   AAW21815;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Cytosine DNA glycosylase.
XX
KW   Cytosine DNA glycosylase; uracil DNA glycosylase; UNG2;
KW   mutagenesis; DNA sequencing; DNA modification; cell killing.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 213
FT                   /note= "site of Asn-204 to Asp substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Claim 6; Refer to page 47-48; 60pp; English.
XX
CC   A novel cytosine DNA glycosylase (CDG) (AAW21815) is capable of
CC   releasing both cytosine and uracil bases from single and/or double
CC   stranded DNA.  It is obtd. by site-directed mutagenesis of human
CC   uracil DNA glycosylase UNG2 cDNA (see also AAT73564) such that the
CC   encoded polypeptide has a Asn-204 to Asp amino acid substn. that
CC   results in gain of CDG activity.  Recombinant enzyme can be
CC   expressed in transformed host cells for use in mutagenesis (in
CC   vivo or in vitro) systems, to remove contaminating DNA prior to
CC   PCR, in DNA modification, in cell killing, and in DNA sequencing
CC   to determine the position of cytokine bases.
XX
SQ   Sequence   313 AA;


  Query Match             63.2%;  Score 1013.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 9.1e-96;
  Matches  187;  Conservative   32;  Mismatches   47;  Indels    3;  Gaps    2;

Qy     35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
          |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db     46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy     93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
          || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||||||||||||||
Db    105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpyhgpnqahg 164

Qy    153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
          ||||||:|||||||| |||||| |||: | ||||||||||||||||:||||||||||
Db    165 lcfsvqrpvpppplseniykelstdiedfvhpghgdlsgwakqgvlldavltvrahqan 224

Qy    213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
          |||:|||| |||||: ||: | |:|||||||| ||| ||||||||  |||||| :
Db    225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy    273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
          ||| ||:|||| | ||: || :||:|: |
Db    285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  10
AAW21824
ID   AAW21824 standard; Protein; 313 AA.
XX
AC   AAW21824;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (N204Q).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 213
FT                   /note= "site of Asn-213 to Gln substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;


  Query Match             63.2%;  Score 1012.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 1.2e-95;
  Matches  187;  Conservative   31;  Mismatches   48;  Indels    3;  Gaps    2;

Qy     35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
```

```
Db     46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy     93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
          || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||||||||||||||
Db    105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpyhgpnqahg 164

Qy    153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
          ||||||:|||||||| |||||| |||: | ||||||||||||||| ||||||||||
Db    165 lcfsvqrpvpppppsleniykelstdiedfvhpghgdlsgwakqgvlllqavltvrahqan 224

Qy    213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
          |||:|||| |||||: ||: |  |:|||||||| |||: |||||||||  |||||| :
Db    225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy    273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
          ||| ||:|||| | ||: || :||:|: |
Db    285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  11
AAW21819
ID   AAW21819 standard; Protein; 313 AA.
XX
AC   AAW21819;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (Q144L).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 153
FT                   /note= "site of Gln-144 to Leu substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;
```

```
  Query Match             63.1%;  Score 1011.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 1.5e-95;
  Matches  187;  Conservative   31;  Mismatches   48;  Indels    3;  Gaps    2;

Qy     35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
          |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db     46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy     93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
          || |||| :|| |||:||  :|||||  ||:: |:||||:||||||||| ||||||||||
Db    105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgldpyhgpnqahg 164

Qy    153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
          ||||||:|||||||| |||||| |||: | |||||||||||||||||||||||||||||
Db    165 lcfsvqrpvpppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy    213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
          |||:|||| ||||||: ||: |  |:|||||||| |||: |||||||||  |||||| :
Db    225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy    273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
          ||| ||:|||| | ||: || :||:|: |
Db    285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  12
AAW21816
ID   AAW21816 standard; Protein; 313 AA.
XX
AC   AAW21816;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Thymine DNA glycosylase.
XX
KW   Thymine DNA glycosylase; uracil DNA glycosylase; UNG2; mutagenesis;
KW   DNA sequencing; DNA modification; cell killing.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 156
FT                   /note= "site of Tyr-147 to Ala substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Claim 13; Refer to page 47-48; 60pp; English.
XX
CC   A novel thymine DNA glycosylase (TDG) (AAW21816) is capable of
CC   releasing both thymine and uracil bases from single and/or double
CC   stranded DNA.  It is obtd. by site-directed mutagenesis of human
CC   uracil DNA glycosylase UNG2 cDNA (see also AAT73564) such that the
CC   encoded polypeptide has a Tyr-147 to Ala amino acid substn. that
CC   results in gain of TDG activity.  Recombinant enzyme can be
CC   expressed in transformed host cells for use in mutagenesis (in
CC   vivo or in vitro) systems, to remove contaminating DNA prior to
```

```
CC   PCR, in DNA modification, in cell killing, and in DNA sequencing
CC   to determine the position of cytokine bases.
XX
SQ   Sequence   313 AA;


  Query Match           63.0%;  Score 1009.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 2.3e-95;
  Matches  187;  Conservative   31;  Mismatches   48;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:|||||||||| ||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpahgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| ||||| ||: |  |:|||||||| |||: |||||||||  |||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtahpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  13
AAW21825
ID   AAW21825 standard; Protein; 313 AA.
XX
AC   AAW21825;
XX
DT   28-SEP-1997  (first entry)
XX
DE   Mutant uracil DNA glycosylase (H268L).
XX
KW   Uracil DNA glycosylase; UNG2; mutagenesis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 277
FT                   /note= "site of His-277 to Leu substn."
XX
PN   WO9725416-A2.
XX
PD   17-JUL-1997.
XX
PF   09-JAN-1997;   97WO-GB00057.
XX
PR   09-JAN-1996;   96GB-0000384.
XX
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (NYFO-) NYFOTEK AS.
XX
PI   Kavli B,  Krokan HE,  Mol CD,  Slupphaug G,  Tainer JA;
XX
DR   WPI; 1997-372857/34.
XX
PT   DNA glycosylase capable of releasing cytosine, thymine or uracil
PT   bases from DNA  - useful in in vitro and/or vivo mutagenesis systems
PT   to remove contaminating DNA prior to PCR amplification
XX
PS   Example 3; Refer to Page 47-48; 60pp; English.
XX
CC   Mutant DNA glycosylases (AAW21819-256) were produced by site-directed
CC   mutagenesis of human uracil DNA glycosylase UNG2 cDNA (see also
CC   AAT73564) and expression in Escherichia coli.  None of these mutants
CC   showed cytosine DNA glycosylase or thymine DNA glycosylase
CC   activity.  In contrast, an N204D substn. (see also AAW21815)
CC   provided cytokine DNA glycosylase activity, and a Y147A substn.
CC   (see also AAW21816) provided thymine DNA glycosylase activity.  The
CC   results demonstrated the significance of Asn204 for specific
CC   binding of uracil-containing DNA and the significance of the Tyr147
CC   side chain ring structure for preventing binding of thymine.
XX
SQ   Sequence   313 AA;


  Query Match           62.9%;  Score 1007.5;  DB 18;  Length 313;
  Best Local Similarity   69.5%;  Pred. No. 3.8e-95;
  Matches  187;  Conservative   31;  Mismatches   48;  Indels    3;  Gaps    2;

Qy      35 ITPKKLRSSNVEQKT--SSPQLSVEQLERMAKNKKAALDKIRAKATPAGFGETWRRELAA 92
           |  ||  :   |  |  ||| || |||:|: :|| ||| :: |:  | ||||:|:: |:
Db      46 ipakkapagqeepgtppssp-lsaeqldriqrnkaaallrlaarnvpvgfgeswkkhlsg 104

Qy      93 EFEKPYFKQLMSFVADERSRHTVYPPADQVYSSTEMCDIQDVKVVILGQDPYHGPNQAHG 152
           || |||| :|| |||:||  :|||||  ||:: |:||||:|||||||||| ||||||||
Db     105 efgkpyfiklmgfvaeerkhytvyppphqvftwtqmcdikdvkvvilgqdpyhgpnqahg 164

Qy     153 LCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAVLTVRAHQAN 212
           ||||||:|||||||| |||||| |||: | ||||||||||||||||||||||||||||||
Db     165 lcfsvqrpvppppsleniykelstdiedfvhpghgdlsgwakqgvlllnavltvrahqan 224

Qy     213 SHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQAVHPSPLSAH 272
           |||:|||| ||||| ||: |  |:|||||||| |||: |||||||||  |||||| :
Db     225 shkergweqftdavvswlnqnsnglvfllwgsyaqkkgsaidrkrhhvlqtalpsplsvy 284

Qy     273 RGFLGCKHFSKANGLLKLSGTEPINWRAL 301
           ||| ||:|||| | ||: || :||:|: |
Db     285 rgffgcrhfsktnellqksgkkpidwkel 313


RESULT  14
AAU69754
ID   AAU69754 standard; Protein; 229 AA.
XX
AC   AAU69754;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Escherichia coli uracil DNA glycosylase (ung) polypeptide.
XX
KW   DNA mutation-binding protein; nuclease; DNA mismatch; cancer;
KW   DNA damage; human xeroderma pigmentosum complementation group; XPF; XPA;
KW   XPC; XPE; ERCC4; human MutS homologue 2; hMSH2; MutS; Nuc; MutY; Fpg;
KW   Fapy-DNA glycosylase; uracil DNA glycosylase; ung; TDG; xthA gene; Uvr A;
KW   A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;
KW   thymine DNA-glycosylase; Uvr B; Uvr C; nth gene; nfo gene; exonuclease;
KW   endonuclease.
XX
OS   Escherichia coli.
XX
PN   WO200173079-A2.
XX
PD   04-OCT-2001.
XX
PF   26-MAR-2001; 2001WO-US09700.
XX
PR   28-MAR-2000; 2000US-192764P.
PR   29-AUG-2000; 2000US-0650855.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Mc Cutchen-maloney SL;
XX
DR   WPI; 2001-656920/75.
DR   N-PSDB; AAS63241.
XX
PT   Recombinant chimeric protein, useful for detecting and quantifying DNA
```

PT mutations, e.g. in disease diagnosis, comprises mutation-binding
PT protein and nuclease -
XX
PS Disclosure; Page 96-97; 128pp; English.
XX
CC Sequences AAU69737-AAU69760 represent proteins which can be used in the
CC synthesis of chimeric proteins comprising a DNA mutation-binding protein,
CC a linker and a nuclease, by recombinant technology. The chimeric proteins
CC are useful for detection, quantification and mapping of DNA sequence
CC variations including mutations, for example, caused by damage and
CC mismatches. The proteins are able to bind to the site of the DNA mutation
CC and cut it out of the molecule. This is useful for early diagnosis of
CC cancer and other diseases. The proteins used in the invention include
CC human XPF (or ERCC4), human xeroderma pigmentosum complementation groups
CC A, C and E (XPA, XPC and XPE), human MutS homologue 2 (hMSH2), Serratia
CC marcescens nuclease (Nuc), Thermus thermophilus MutS, Escherichia coli
CC Fapy-DNA glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific
CC adenine glycosylase (MutY), synthetic T4 endonuclease V (T4 endo V),
CC thymine DNA-glycosylase (TDG), E. coli Uvr A, B and C, and E. coli
CC endonucleases and exonucleases.
XX
SQ Sequence 229 AA;

Query Match 42.5%; Score 681.5; DB 22; Length 229;
Best Local Similarity 60.9%; Pred. No. 9e-62;
Matches 131; Conservative 24; Mismatches 59; Indels 1; Gaps 1;

```
Qy      85 TWRRELAAEFEKPYFKQLMSFVADER-SRHTVYPPADQVYSSTEMCDIQDVKVVILGQDP 143
           ||   || | ::|||   :  || || |  |:|||   |:::    :: ||||||||||
Db       6 twhdvlaeekqqpyflntlqtvaserqsgvtiyppqkdvfnafrftelgdvkvvilgqdp 65

Qy     144 YHGPNQAHGLCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNAV 203
           |||| ||||| |||:  :  ||||:|:||||   | ||  | || |  ||:||||||| |
Db      66 yhgpgqahglafsvrpgiaippsllnmykelentipgftrpnhgyleswarqgvlllntv 125

Qy     204 LTVRAHQANSHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQA 263
           ||||| ||:||   |||||||| ||  :: :|||||||||||:| |||| ||::||||||:|
Db     126 ltvragqahshaslgwetftdkvislinqhregvvfllwgshaqkkgaiidkqrhhvlka 185

Qy     264 VHPSPLSAHRGFLGCKHFSKANGLLKLSGTEPINW 298
           ||||||||||| || ||  ||  |:  |  ||:|
Db     186 phpsplsahrgffgcnhfvlanqwleqrgetpidw 220
```

RESULT 15
AAB88498
ID AAB88498 standard; Protein; 219 AA.
XX
AC AAB88498;
XX
DT 04-JUN-2001 (first entry)
XX
DE Haemophilus influenzae essential bacterial protein SEQ ID NO:14.
XX
KW Haemophilus influenzae; essential bacterial gene; identification;
KW otitis media; meningitis; upper respiratory tract infection;
KW infection; antimicrobial.
XX
OS Haemophilus influenzae.
XX
PN WO200111033-A2.
XX
PD 15-FEB-2001.
XX
PF 03-AUG-2000; 2000WO-US21176.
XX
PR 04-AUG-1999; 99US-0368382.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Chovan LE, Hessler PE, Reich KA;
XX
DR WPI; 2001-147511/15.
DR N-PSDB; AAF94351.
XX
PT Essential bacterial genes from Haemophilus influenzae and methods for
PT identifying 'essential' genes that may be potential therapeutic targets
PT -
XX
PS Claim 9; Page 53; 185pp; English.
XX
CC AAF94345 to AAF94409 represent essential bacterial genes from
CC Haemophilus influenzae, which encode the proteins given in AAB88492 to
CC AAB88556. The present invention also describes methods for identifying
CC essential bacterial genes (i.e. those essential to the survival of a
CC bacterium) using a transposition system. The methods are used to
CC identify essential genes from bacteria, especially H. influenzae (which
CC causes otitis media, meningitis and upper respiratory tract infections)
CC which may be used as targets for potential antimicrobial agents.
CC AAF94410 to AAF94416 represent PCR primers used in the exemplification
CC of the present invention.
XX
SQ Sequence 219 AA;

Query Match 39.6%; Score 635.5; DB 22; Length 219;
Best Local Similarity 57.1%; Pred. No. 4.6e-57;
Matches 124; Conservative 28; Mismatches 64; Indels 1; Gaps 1;

```
Qy      84 ETWRRELAAEFEKPYFKQLMSFVADER-SRHTVYPPADQVYSSTEMCDIQDVKVVILGQD 142
           : |   :  |  :|||:  :  |   | |  |:||| : |::: :    :||||||||||
Db       2 knwtdvigtekaqpyfqhtlqqvhlarasgktiyppqedvfnafkytafedvkvvilgqd 61

Qy     143 PYHGPNQAHGLCFSVQKPVPPPPSLVNIYKELCTDIDGFKHPGHGDLSGWAKQGVLLLNA 202
           |||||||||||| |||:  |  ||||:||||||  || ||: | :| |  ||:|||||||
Db      62 pyhgpnqahglafsvkpevaippsllniykeltqdisgfqmpsngylvkwaeqgvlllnt 121

Qy     203 VLTVRAHQANSHKDRGWETFTDAVIKWLSVNREGVVFLLWGSYAHKKGATIDRKRHHVLQ 262
           ||||    |:|| : ||| ||| ||  |: :|| :||||||:| |||  ||| || ||
Db     122 vltvergmahshanlgwerftdkviavlnehreklvfllwgshaqkkgqmidrtrhlvlt 181

Qy     263 AVHPSPLSAHRGFLGCKHFSKANGLLKLSGTEPINWR 299
           | |||||||||| ||:|||| |  |:  | :||:|:
Db     182 aphpsplsahrgffgcrhfsktnsyleshgikpidwq 218
```

Search completed: August 21, 2002, 10:24:34
Job time: 551 sec